Current Biology 18, 191–194, February 12, 2008 ©2008 Elsevier Ltd All rights reserved DOI 10.1016/j.cub.2007.12.053

Report

The Representation of Egomotion in the Human Brain

Matthew B. Wall[1] and Andrew T. Smith[1,*]
[1]Department of Psychology
Royal Holloway, University of London
Egham, TW20 0EX
United Kingdom

Summary

An essential function of visual processing is to establish the position of the body in space and, in concert with the other sense systems, to monitor movement of the whole body, or "egomotion." A key cue to egomotion is optic flow. For example, forward motion through the environment generates an expanding pattern of flow on the retina, and (with eyes fixed centrally) the direction of heading corresponds to the center of expansion [1]. In macaques, visual cortical area MST is sensitive to optic-flow structure [2, 3], and it has been suggested that MST has a central role in the computation of heading [4]. However, here we identify two areas of the human brain that represent visual cues to egomotion more directly than does MST. These areas respond strongly to a single optic-flow stimulus but become relatively unresponsive when the stimulus is surrounded with further flow patches and thereby made inconsistent with egomotion. One is putative area VIP in the anterior portion of the intraparietal sulcus. The other is a new visual area, which we refer to as cingulate sulcus visual area (CSv). Areas V1–V4 and MT respond about equally to both types of flow stimulus. MST has intermediate properties, responding well to multiple patches but with a modest preference for a single, egomotion-compatible patch. We suggest that MST is merely an intermediate processing stage for visual cues to egomotion and that such cues are more comprehensively encoded by VIP and CSv.

Results

Many cortical regions, from V1 upward, respond well to an expanding pattern of dots, but this does not imply the encoding of egomotion. Using fMRI, we have applied a simple, strong test for specificity of responses to egomotion-compatible visual stimuli. We presented moving-dot arrays that contained either a single patch or multiple patches of optic flow (Figures 1A and 1B, respectively). Any cortical areas containing neurons that are responsive to the presence of global flow structure irrespective of context, or indeed are responsive to local dot motion, should respond to both stimuli. In contrast, a brain region that is active only when retinal stimulation is indicative of egomotion should only respond to the single stimulus, because a flow pattern generated by egomotion can have only one center of flow. To maximize responses, we used time-varying optic flow, consistent (for a single patch) with back-and-forth, spiraling egomotion [5]. Six participants were scanned at 3T with conventional acquisition procedures. They performed a counting task at a central fixation point to engage attention and ensure good fixation. Video eye-tracking data showed that eye-position variance was acceptable (mean standard deviation [SD] = 0.58 deg) and did not vary significantly among the various stimuli used. Standard retinotopic mapping was used to define the boundaries of visual areas V1–V4. The MT+ complex was identified and MST was distinguished from MT by use of the criterion of sensitivity to ipsilateral stimulation [6]. Figure 1C shows response magnitudes for single and multiple motion patches in each occipital visual area studied. In V1, V2, V3, V3A, V3B, and V4, an array of nine optic-flow patches elicited a response that was as strong as or stronger than that produced by a single, large patch containing the same total number of dots, each with the same size and speed. Area MST, previously associated with flow specificity in both nonhuman primates [2] and humans [7], showed a subtly different result (Figure 1D, red broken line). Here, the response to nine patches was significantly weaker than the response to one [$t(5) = 2.91$, $p < 0.05$]. This is consistent with the presence of neurons tuned to visual motion that is compatible with egomotion. However, it is far from the case that MST was silent in the presence of incompatible motion; the reduction was only about 15%. The results for MT (Figure 1D, red solid line) differed little from those for V1–V4, although there was a trend toward the behavior of MST in comparison with those areas.

Having obtained results for this set of established, independently defined visual areas, we searched for additional brain regions that might be differentially sensitive to a single flow pattern by using a statistical comparison of responses to the one-patch and nine-patch stimuli. Using standard methods, we conducted a voxel-wise search and found two clusters (other than MST) that were more responsive to one patch than to nine. One was present in ten of 12 hemispheres and was consistently located in the fundus of the anterior intraparietal sulcus. The mean coordinates in Talairach space (L: −26 −53 40; R: 26 −48 40) are close to those of polysensory motion-sensitive areas documented previously [8, 9] and to those of visual areas IPS2 and IPS3 [10]. Pending allocation to one of the subregions in this vicinity, we refer to our focus simply as putative VIP, in line with [8, 9], although it is not clear that our VIP is the same as that identified in either of those studies or that the two studies identify the same area as each other. The second cluster was at the boundary of the medial frontal cortex and the limbic lobe, in the cingulate sulcus, and was clearly identifiable in nine hemispheres (average coordinates, L: −10 −25 38; R: 10 −26 41). This region is not usually associated with optic flow, but the nearby posterior cingulate cortex has been associated with visual navigation in both macaques [11] and humans [12], so it is perhaps a natural candidate for involvement in the representation of egomotion. We provisionally refer to this new visual area as cingulate sulcus visual area (CSv).

Figure 2 shows the locations of CSv and VIP, together with the associated response magnitudes from our experiment. Both areas show a strong response to a single patch, but this response is severely reduced when the stimulus is rendered inconsistent with egomotion by the inclusion of additional centers of motion. The reduction was quantified as an attenuation index $[(R_1 - R_9) / R_1]$ where R_1 and R_9 are the normalized responses to the one-patch and nine-patch stimuli. The mean

*Correspondence: a.t.smith@rhul.ac.uk

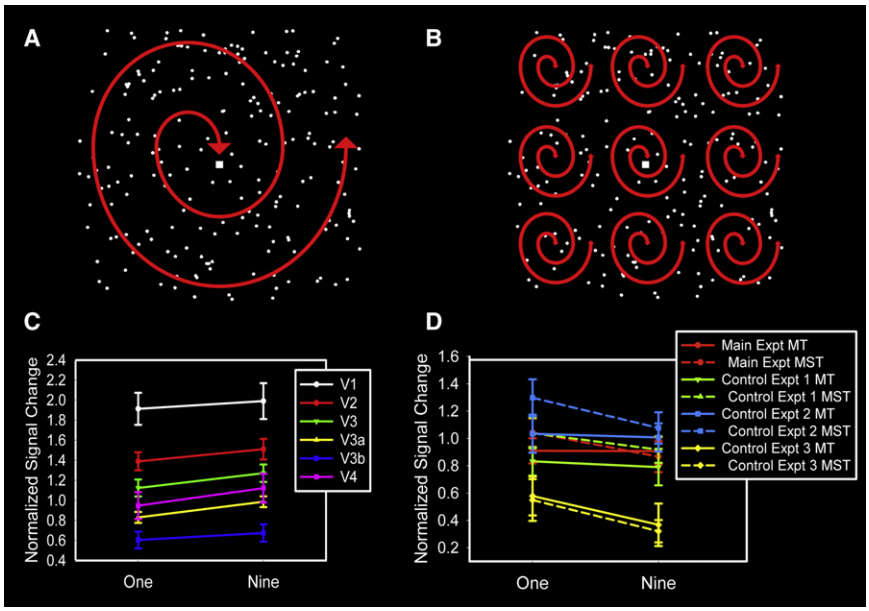

Figure 1. Stimuli and Results for the Retinotopic Visual Areas

(A and B) Diagrammatic representation of the motion stimuli used. The stimulus was either a single patch of time-varying optic flow (A) or an array of nine similar patches (B).

(C and D) Mean activation (12 hemispheres) in the retinotopic areas V1–V4 (C, main experiment only) and in MT and MST (D, main experiment and three control experiments; see text for details). Activation is shown normalized with respect to the mean activity across all areas and conditions in each subject (mean = 1.0). Error bars represent ± 1 standard error.

attenuation was 46% in VIP and 77% in CSv. In both cases, this is substantially greater than the 15% attenuation found in MST. The difference between MST and VIP was significant [$t(5) = 3.0$, $p < 0.03$], as were the differences between VIP and CSv [$t(5) = 5.4$, $p < 0.003$] and between MST and CSv [$t(5) = 3.0$, $p < 0.001$]. To avoid biasing the result by taking measurements only from voxels that showed a significant difference between the two stimuli, we redefined each of the two regions by using an independent statistical comparison, based on a separate experiment in which activity obtained in the single-patch condition was compared to that measured during fixation with no stimulus. This revealed many active regions, as expected. From these, a voxel cluster was identified at each of the two locations described above, and this was used to compute the mean activations shown in Figure 2. Such clusters were identified in 11 of 12 hemispheres for VIP and in 12 of 12 for CSv; all were included whether or not similar clusters had been apparent in the original contrast between the two stimuli.

Similar results were obtained in three control experiments. In control experiment 1, each of the nine stimulus patches was a scaled version of the single patch (i.e., moving dots were proportionately smaller, denser, and slower than those in the single patch). This controls for the possibility that the local dot properties in the main experiment might be less appropriate for small patches than for large ones. In control experiment 2, the entire display was made much smaller. This controls for the possibility that responses might be lost in the original nine-patch condition simply because each flow pattern is too small to be effective; if this were true then a single small patch would be similarly affected in the control experiment. The results of these experiments are included in Figure 1D for MT and MST and in Figure 2 for VIP and CSv (see Supplemental Data, available online, for V1–V4). They are very similar, showing that these factors do not explain the results. In control experiment 3, the motion-defined edges in the nine-patch stimulus were severely blurred by randomizing the location of disappearance of dots reaching the boundaries, within a range of ± 1.1° of the nominal boundary location. This was done to test the possibility that motion-defined contours might inhibit VIP and CSv, causing the observed reduction for multiple patches. In addition, the time-varying flow was replaced by continuous expansion, as a check that this more commonly used stimulus gives the same result. Again, the results for the key areas are included in Figures 1 and 2, and the remainder are shown in the Supplemental Data. Activations are markedly reduced for one-patch expansion compared with time-varying flow, as we have previously found for MT and MST [7], but the pattern of results is similar, the nine-patch stimulus giving no measurable activation in either VIP or CSv.

Discussion

In macaques, optic-flow components such as expansion and rotation are represented explicitly in area MSTd in the posterior temporal cortex. Some neurons in this area are sensitive to the location of the focus of expansion [3, 13], and it has been suggested that they signal direction of heading. This has led to the formulation of models of heading and navigation based on MSTd response properties (e.g., [14]). MSTd also has vestibular inputs, and it has been suggested that these feed into heading perception at the level of MSTd [15]. However, vestibular-direction tuning is often opposite to that for visual stimuli (particularly for rotation), suggesting that one purpose of vestibular input may be to compensate for egomotion rather than to assist in encoding it [16]. Human MST has been identified [6] and shown to be sensitive to optic-flow characteristics [7]. We have no reason to doubt that MST is involved in encoding heading, in both monkeys and humans, but we show here that strong activity can occur in human MST in response to visual stimuli that are inconsistent with egomotion. This makes it unlikely that MST signals egomotion directly, although it is possible that only a minority of MST cells encode heading and that our results reflect the activity of the remainder.

Macaque MSTd projects to posterior parietal areas including polysensory area VIP in the ventral intraparietal sulcus. Here, neurons have visual properties similar to MSTd [17] but are more sensitive than MSTd neurons to motion in the vestibular and somatosensory modalities [18, 19] and many response fields are in craniocentric coordinates [20]. It has been suggested [20] that primate VIP is a more likely substrate of egomotion than MST is, but this has remained speculative. In humans, VIP has been identified and shown to be motion-sensitive in multiple sensory modalities [8, 9]. We interpret our findings as evidence that human VIP may have a more central role than MST in extracting visual cues to egomotion. It is likely that VIP receives input from MST, as in macaques, and that MST is

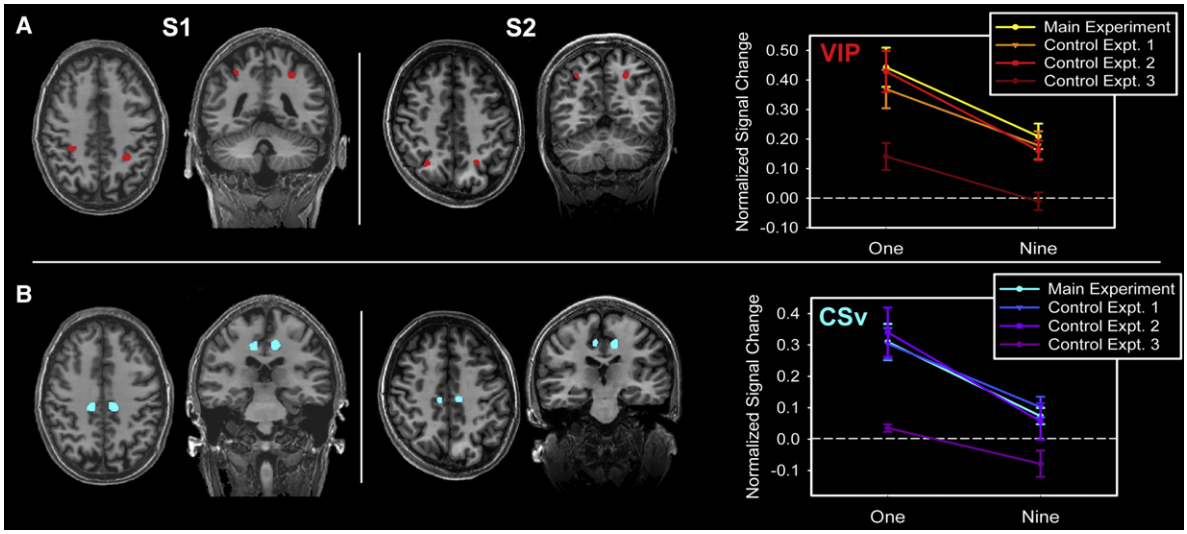

Figure 2. Results for VIP and CSv

The locations, in two representative participants (labeled S1 and S2), of two cortical regions that respond preferentially to a single flow pattern, together with associated mean response amplitudes averaged across all hemispheres, for all experiments (normalized, as in Figure 1) (A) shows putative area VIP (ventral intraparietal). (B) shows a new visual area, CSv (cingulate sulcus visual area). Both areas respond strongly to a single stimulus, but their responses are severely reduced in the presence of multiple flow patches. Error bars represent ±1 standard error.

thus an intermediate processing stage in a pathway that culminates in VIP (or even beyond). From there, information about heading could readily feed into motor-control systems. We have no information about whether our VIP is craniotopically organized, but some evidence of craniotopic receptive fields exists even in the MT complex in humans [21, 22] as well as in macaques [23].

The role of CSv is more difficult to evaluate. It is possible that this area is the homolog of macaque posterior cingulate cortex (PCC), which has been shown to contain neurons sensitive to visual stimuli and to saccades [11, 24] within an allocentric coordinate system [25]. But these studies suggest a role in detecting salient events for use in spatial orienting of attention, rather than for navigation. Alternatively, it is possible that CSv, which is in the depths of the cingulate sulcus and not in the cingulate gyrus, is not homologous with PCC but corresponds to an imprecisely defined area in the vicinity of the human cingulate cortex that has been shown to be concerned with navigation [12, 26]. For example, heading information might be received in CSv from VIP and used in relation to complex navigation tasks. Finally, we cannot rule out a central role for CSv in sensory encoding of heading, but the macaque literature makes it more likely that this function is associated with VIP.

In summary, we conclude that beyond MST at least two brain areas exist that respond well to optic flow only when it is consistent with egomotion. One (VIP) is in a known complex of visual areas, and the other (CSv) is a new visual area. One or both of these areas may directly encode visual cues to egomotion.

Experimental Procedures

Stimuli and Procedure

All stimuli consisted of high-contrast, moving, random dot patterns (light dots of approximate luminance 700 cd/m^2 on a dark background). Each dot moved in a straight path at a speed of 10°/s for a lifetime of 133 ms (ten frames) before disappearing and reappearing at a new, random location. Dots leaving the edge of the stimulus or reaching the center of the spiral disappeared for the remainder of their lifetime before being replotted. Different dots were repositioned at different times, 10% of the dots being repositioned at each frame update. Global patterns of optic flow were produced by control of the local motion directions of the dots. A fixation point was continuously present.

The main experiment contained two stimulus conditions. The first consisted of a 20° × 20° square field of 800 dots moving in a coherent optic-flow pattern containing expansion/contraction and rotation components that varied over time, consistent with self-motion on a varying spiral trajectory. This stimulus is derived from [5]. The second stimulus condition consisted of the same 20° × 20° stimulus area divided into nine identical panels, each containing a moving stimulus similar to condition 1 but smaller. Dots were removed and randomly replotted if they reached the edge of any panel, giving a sharp motion-defined boundary. The dot size, dot speed, and number of dots in the whole array were made identical across conditions in order to equate low-level visual characteristics.

Stimuli were presented for periods of 5 s, within an event-related fMRI paradigm. The intertrial intervals (ITIs) were determined by a Poisson distribution [27] with a mean of 5.5 s and a range of 2–10 s. There were 32 trials in each scanning run, and each run had a duration of 5 min 50 s (including a 10 s buffer period at the beginning). There were six participants, and each completed six such scanning runs for each of the four experiments (main experiment and three controls). The stimulus sequence and ITI sequence were determined pseudorandomly for each run. All participants were presented with the same six sequences but in different random orders. Throughout each run, participants completed a simple counting task at the fixation point in order to ensure constant maintenance of fixation. The fixation point changed color randomly at a rate of 2 Hz; participants were instructed to count the number of blue fixation points occurring throughout each scanning run and to report the total verbally at the end of each run.

Scanning runs to define regions of interest (ROIs) were usually performed in a separate session. MT and MST were defined by the use of an ipsilateral stimulus based on the method used in [6] and [28] and also previously used in our lab [7]. Retinotopic areas V1–V4 were identified by a standard retinotopic mapping procedure [29], with a counterphasing checkerboard "wedge" stimulus (a 24° sector) of radius 12°. Further details are provided in the Supplemental Data.

A third localizer stimulus was used in order to identify any areas that are outside the occipital lobe and might also be important. This consisted of a coherent-flow stimulus, identical to the single-patch stimulus used in the main experiment and described above, which alternated with a blank (except for a fixation point) screen in blocks of 15 s. This 30 s stimulus sequence was repeated ten times to give a run length of 310 s (including a 10 s buffer period at the beginning), and participants completed two such scanning runs.

Data Acquisition and Analysis

In brief, images were acquired with a 3T MR scanner via conventional procedures and 3 mm isotropic voxels. Data were preprocessed and analyzed with BrainVoyager QX. All data (except the retinotopic mapping data, see below) were analyzed by the application of the General Linear Model. MST was defined as all contiguous voxels within the MT+ complex that were significantly active during ipsilateral-motion stimulation. MT was defined as all contiguous voxels that were active during contralateral but not ipsilateral stimulation, excluding any voxels situated further anterior than the median value of the MST ROI. Retinotopic data were analyzed conventionally in terms of temporal phase [29], and ROIs for visual areas V1–V4 were defined on a flattened version of each participant's reference anatomy. Data from the third localizer were thresholded at $p < 0.001$ (uncorrected) in order to identify any areas outside the occipital lobe that were sensitive to visual motion. From the various active areas, two clusters were selected that corresponded to the locations of the clusters consistently seen when the one-patch and nine-patch conditions from the main experimental data were contrasted. For the main experimental data, ROI-based analyses were conducted in order to extract the mean response amplitude related to each condition for each ROI.

Supplemental Data

Additional Experimental Procedures and one figure can be found online at http://www.current-biology.com/cgi/content/full/18/3/191/DC1/.

Acknowledgments

We thank Jonas Larsson, John Wann, and three anonymous reviewers for their helpful comments on an earlier version.

Received: October 16, 2007
Revised: December 21, 2007
Accepted: December 21, 2007
Published online: January 24, 2008

References

1. Gibson, J.J. (1950). The Perception of the Visual World (Boston: Houghton Mifflin).
2. Saito, H.A., Yukie, M., Tanaka, K., Hikosaka, K., Fukada, Y., and Iwai, E. (1986). Integration of direction signals of image motion in the superior temporal sulcus of the macaque monkey. J. Neurosci. *6*, 145–157.
3. Tanaka, K., and Saito, H. (1989). Analysis of motion of the visual field by direction, expansion/contraction, and rotation cells clustered in the dorsal part of the medial superior temporal area of the macaque monkey. J. Neurophysiol. *62*, 626–641.
4. Duffy, C.J., and Wurtz, R.H. (1995). Responses of monkey MST neurons to optic flow stimuli with shifted centers of motion. J. Neurosci. *15*, 5192–5208.
5. Morrone, M.C., Tosetti, M., Montanaro, D., Fiorentini, A., Cioni, G., and Burr, D.C. (2000). A cortical area that responds specifically to optic flow, revealed by fMRI. Nat. Neurosci. *3*, 1322–1328.
6. Dukelow, S.P., DeSouza, J.F.X., Culham, J.C., van den Berg, A.V., Menon, R.S., and Vilis, T. (2001). Distinguishing subregions of the human MT+ complex using visual fields and pursuit eye movements. J. Neurophysiol. *86*, 1991–2000.
7. Smith, A.T., Wall, M.B., Williams, A.L., and Singh, K.D. (2006). Sensitivity to optic flow in human cortical areas MT and MST. Eur. J. Neurosci. *23*, 561–569.
8. Bremmer, F., Schlack, A., Shah, N., Zafiris, O., Kubischik, M., Hoffmann, K.-P., Zillies, K., and Fink, G. (2001). Polymodal motion processing in posterior parietal and premotor cortex: A human fMRI study strongly implies equivalencies between humans and monkeys. Neuron *29*, 287–296.
9. Sereno, M.I., and Huang, R.-S. (2006). A human parietal face area contains aligned head-centered visual and tactile maps. Nat. Neurosci. *9*, 1337–1343.
10. Swisher, J.D., Halko, M.A., Merabet, L.B., McMains, S.A., and Somers, D.C. (2007). Visual topography of human intraparietal sulcus. J. Neurosci. *27*, 5326–5337.
11. Dean, H.L., Crowley, J.C., and Platt, M.L. (2004). Visual and saccade-related activity in macaque posterior cingulate cortex. J. Neurophysiol. *92*, 3056–3068.
12. Grön, G., Wunderlich, A.P., Spitzer, M., Tomczak, R., and Riepe, M.W. (2000). Brain activation during human navigation: Gender-different neural networks as substrate of performance. Nat. Neurosci. *3*, 404–408.
13. Duffy, C.J., and Wurtz, R.H. (1991). Sensitivity of MST neurons to optic flow stimuli. I. A continuum of response selectivity to large-field stimuli. J. Neurophysiol. *65*, 1329–1345.
14. Grossberg, S., Mingolla, E., and Pack, C. (1999). A neural model of motion processing and visual navigation by cortical area MST. Cereb. Cortex *9*, 878–895.
15. Gu, Y., DeAngelis, G.C., and Angelaki, D.E. (2007). A functional link between area MSTd and heading perception based on vestibular signals. Nat. Neurosci. *10*, 1038–1047.
16. Takahashi, K., Gu, Y., May, P.J., Newlands, S.D., DeAngelis, G.C., and Angelaki, D.E. (2007). Multimodal Coding of Three-Dimensional Rotation and Translation in Area MSTd: Comparison of Visual and Vestibular Selectivity. J. Neurosci. *27*, 9742–9768.
17. Schaafsma, S.J., and Duysens, J. (1996). Neurons in the ventral intraparietal area of awake macaque monkey closely resemble neurons in the dorsal part of the medial superior temporal area in their responses to optic flow patterns. J. Neurophysiol. *76*, 4056–4068.
18. Duhamel, J.R., Colby, C.L., and Goldberg, M.E. (1998). Ventral intraparietal area of the macaque: Congruent visual and somatic response properties. J. Neurophysiol. *79*, 126–136.
19. Bremmer, F., Klam, F., Duhamel, J.R., Hamed, S.B., and Graf, W. (2002). Visual-vestibular interactive responses in the macaque ventral intraparietal area (VIP). Eur. J. Neurosci. *16*, 1569–1586.
20. Zhang, T., Heuer, H.W., and Britten, K.H. (2004). Parietal area VIP neuronal responses to heading stimuli are encoded in head-centered coordinates. Neuron *42*, 993–1001.
21. Goossens, J., Dukelow, S.P., Menon, R.S., Vilis, T., and van den Berg, A.V. (2006). Representation of head-centric flow in the human motion complex. J. Neurosci. *26*, 5616–5627.
22. d'Avossa, G., Tosetti, M., Crespi, S., Biagi, L., Burr, D.C., and Morrone, M.C. (2007). Spatiotopic selectivity of BOLD responses to visual motion in human area MT. Nat. Neurosci. *10*, 249–255.
23. Bradley, D.C., Maxwell, M., Andersen, R.A., Banks, M.S., and Shenoy, K.V. (1996). Mechanisms of heading perception in primate visual cortex. Science *273*, 1544–1547.
24. Olson, C., Musil, S., and Goldberg, M. (1993). Posterior cingulate cortex and visuospatial cognition: Properties of single neurons in the behaving monkey. In Neurobilogy of Cingulate Cortex and Limbic Thalamus: A Comprehesive Handbook, B.a.G, M. Vogt, ed. (Boston: Birkhauser), pp. 366–380.
25. Dean, H.L., and Platt, M.L. (2006). Allocentric spatial referencing of neuronal activity in macaque posterior cingulate cortex. J. Neurosci. *26*, 1117–1127.
26. Pine, D.S., Grun, J., Maguire, E.A., Burgess, N., Zarahn, E., Koda, V., Fyer, A., Szeszko, P.R., and Bilder, R.M. (2002). Neurodevelopmental aspects of spatial navigation: A virtual reality fMRI study. Neuroimage *15*, 396–406.
27. Hagberg, G., Zito, G., Patria, F., and Sanes, J. (2001). Improved detection of event-related functional MRI signals using probability functions. Neuroimage *14*, 1193–1205.
28. Huk, A.C., Dougherty, R.F., and Heeger, D.J. (2002). Retinotopy and functional subdivision of human areas MT and MST. J. Neurosci. *22*, 7195–7205.
29. Sereno, M.I., Dale, A.M., Reppas, J.B., Kwong, K.K., Belliveau, J.W., Brady, T.J., Rosen, B.R., and Tootell, R.B.H. (1995). Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. Science *268*, 889–893.